

# Dokumentation und Annotationsrichtlinien für das Korpus BeDiaCo Version 1

Malte Belz, Alina Zöllner, Megumi Terada

Institut für deutsche Sprache und Linguistik  
Humboldt-Universität zu Berlin

15. April 2020

DOI: <https://doi.org/10.18452/21361>

# Inhaltsverzeichnis

<b>1</b>	<b>Korpus</b>	<b>4</b>
1.1	Zusammenfassung und Beschreibung	4
1.2	Experimentablauf	5
1.3	Material	6
1.3.1	Freies Gespräch	6
1.3.2	Diapixe	6
1.3.3	Wortlisten	6
1.4	Dateibenennung	7
1.5	Audionachverarbeitung und Pseudonymisierung	7
1.6	Überblick über die Annotationen	8
1.7	Verfügbarkeit	8
1.8	Datenstruktur im Medienrepositorium	9
1.9	Versuchspersonenmetadaten und Dyadenzuordnung	9
1.10	Einwilligungserklärungen	10
<b>2</b>	<b>Annotationsschema</b>	<b>13</b>
2.1	Aufgabenfreie Dialoge	13
2.2	Wortlisten	13
<b>3</b>	<b>Annotationsebenen</b>	<b>14</b>
3.1	Ebene <b>dip1</b> (Transliteration)	14
3.1.1	Erstellung	14
3.1.2	Annotationswerte	16
3.2	Ebene <b>fp</b> (Füllpartikeln)	20
3.2.1	Erstellung	20
3.2.2	Annotationswerte	20
3.3	Ebene <b>segm</b> (Segmentale Annotation von Füllpartikeln)	21
3.3.1	Erstellung	22
3.3.2	Annotationswerte	22
3.4	Ebene <b>phon</b> (Phonationsartannotation der Segmente)	23
3.4.1	Erstellung	23
3.4.2	Annotationswerte	23
3.5	Ebene <b>ip</b> (Intonationsphrasenannotation)	23
3.5.1	Erstellung	23
3.5.2	Annotationswerte	24
3.6	Ebene <b>dia</b> (Dialogzügeannotation in BeDiaCo)	24
3.6.1	Erstellung	24
3.6.2	Annotationswerte	24

3.7	Ebene <b>pseudo</b> (Anonymisierung)	25
3.7.1	Erstellung	25
3.7.2	Annotationswerte	26
3.8	Ebene <b>ORT</b>	26
3.8.1	Erstellung	26
3.8.2	Annotationswerte	26
3.9	Ebene <b>MAU-cor</b>	27
3.9.1	Erstellung	27
3.9.2	Annotationswerte	27
3.10	Ebene <b>vowel</b>	27
3.10.1	Erstellung	27
3.10.2	Annotationswerte	27
<b>4</b>	<b>Emu-Datenbank</b>	<b>28</b>
4.1	Erstellen der emuR-Datenbank	28
4.2	Enthaltene Signalebenen und Korrekturen	31
4.3	Darstellung im Browser	31
	<b>Literatur</b>	<b>32</b>

# 1 Korpus

## 1.1 Zusammenfassung und Beschreibung

Tabelle 1 enthält die Basisdaten von BeDiaCo v. 1.

**Tabelle 1:** Zusammenfassung von BeDiaCo v. 1.

Name	Berlin Dialogue Corpus
Abkürzung	BeDiaCo
Version	1
Herausgebende	Malte Belz, M. A.
Adresse	Prof. Dr. Christine Mooshammer Institut für deutsche Sprache und Linguistik Humboldt-Universität zu Berlin Unter den Linden 6 10099 Berlin
Zitation des Korpus	Malte Belz & Christine Mooshammer (2020): <i>Berlin Dialogue Corpus (BeDiaCo). Version 1</i> . Humboldt-Universität zu Berlin. URL: <a href="https://rs.cms.hu-berlin.de/phon">https://rs.cms.hu-berlin.de/phon</a>
Zitation dieser Dokumentation	Malte Belz et al. (2020): <i>Dokumentation und Annotationsrichtlinien für das Korpus BeDiaCo Version 1</i> . Humboldt-Universität zu Berlin. DOI: <a href="https://doi.org/10.18452/21361">10.18452/21361</a>
Zugang	Medienrepositorium der HU, <a href="https://rs.cms.hu-berlin.de/phon/">https://rs.cms.hu-berlin.de/phon/</a>
Annotator*innen	Malte Belz, Alina Zöllner, Megumi Terada
Dialoge	8
Versuchspersonen	10 m, 6 f
Sprache	Deutsch
Register	Aufgabenfreie spontansprachliche Dialoge
Zusatzmaterial	Wortlisten
Annotationen	Diplomatische Ebene der Dialoge ( <b>dipl</b> ) Füllpartikeln ( <b>fp</b> ) Segmente ( <b>segm</b> ) Phonation ( <b>phon</b> ) Dialogstruktur ( <b>dia</b> ) Intonationsphrasen ( <b>ip</b> ) Pseudonymisierung ( <b>pseudo</b> ) Wortlistentransliteration ( <b>ORT</b> ) Korrektur der automatischen Segmentierung aus ORT ( <b>MAU-cor</b> ) Vokale in Wortlisten ( <b>vowel</b> )

Das Korpus BeDiaCo v. 1 (Berlin Dialogue Corpus v.1) besteht aus akustischen Aufnahmen spontansprachlicher Dialoge deutscher Muttersprachler\*innen mit sowohl aufgabenfreien als auch aufgabenbasierten Teilen und zusätzlichen gelesenen Wortlisten. BeDiaCo v. 1 wurde im Jahr 2018 und 2019 im Rahmen einer Disser-

tation erhoben (Belz [in Vorbereitung]) und vom Institut für deutsche Sprache und Linguistik mit 290 Euro für die Versuchspersonenhonorare gefördert.<sup>1</sup> Für BeDiaCo v. 1 wurden ursprünglich 28 Versuchspersonen (14 Dyaden) aufgenommen, in eine Weiternutzung durch akademische Dritte haben jedoch nur 16 Versuchspersonen (8 Dyaden) eingewilligt.

Tabelle 2 enthält die Anzahl der Gesamttoken, der bereinigten Token (ohne Token mit spitzen Klammern, wie beispielsweise Pausen, unverständliche Ausdrücke, etc., und Token mit Leerzeichen), Silben, Artikulationszeit (min) und Artikulationsgeschwindigkeit in Silben je Sekunde ( $\sigma/s$ ).

**Tabelle 2:** Anzahl der Gesamttoken (Tok. ges.), bereinigten Token (Tok. part.), sowie der auf die partiellen Token bezogene Silbenanzahl, Artikulationsdauer (min) und Artikulationsgeschwindigkeit in Silben je Sekunde ( $\sigma/s$ )

VP	Tok. ges.	Tok. part.	Artikulationsdauer	Silben	$\sigma/s$
f1	1750	1402	6,25	1943	5,18
f2	2268	1888	8,51	2661	5,21
f7	1875	1534	6,38	2238	5,85
f10	2492	2146	8,17	2971	6,06
f12	1616	1338	5,32	1860	5,83
f13	1174	964	4,28	1334	5,19
m1	1916	1638	8,90	2510	4,70
m4	2041	1724	7,17	2442	5,68
m5	1192	988	4,45	1385	5,19
m6	1996	1660	8,08	2424	5,00
m7	1506	1181	5,73	1695	4,93
m8	1505	1211	6,30	1774	4,69
m10	2066	1824	7,27	2668	6,12
m13	2247	1909	8,31	2947	5,91
m14	1416	1207	4,08	1730	7,07
m15	2293	1957	9,94	2898	4,86
$\Sigma$	29353	24571	109,14	35480	5,42
$\bar{x}$			12,84	4174	5,46

Tabelle 2 wird dynamisch mit dem R-Paket *knitr* v. 1.28 erzeugt und ist der Dokumentation zur Replizierbarkeit beigelegt (*loadandquery.Rnw*).

## 1.2 Experimentablauf

Die Aufnahme erfolgte mit zwei Nackenbügelsprechgarnituren von beyerdynamics (Headset Opus 54), die an einen XLR-Adapter mit den Einstellungen ‚-‘ (keine Filterung tiefer und hoher Frequenzen) und ‚0‘ (keine Reduzierung der Lautstärke um

<sup>1</sup>Das vorliegende Korpus ist trotz ähnlicher Namensgebung unabhängig von den in Rasskazova et al. (2014) beschriebenen Daten.

12 dB) angeschlossen sind. Das Audiosignal lief in einen externen Verstärker (*tascam*) mit +48 V Phantomspannung und zwei Kanälen. Zu Beginn stand die Verstärkung beider Kanäle auf 10 Uhr und wurde an die Sprechlautstärke der Versuchspersonen angepasst.

Die Aufnahmen wurden im Phonetiklabor des Instituts für deutsche Sprache und Linguistik durchgeführt. Das Experiment dauerte ca. eine Stunde, die VP erhielten ein Honorar von 10 Euro. Der Ablauf war folgendermaßen festgelegt.

- Ankommen, Ausfüllen der Einwilligung (15 min)
- Verkabelung, Testen, VP 1 Wortliste 1 (5 min)
- Verkabelung, Testen, VP 2 Wortliste 1 (ohne VP 1) (5 min)
- Diapix 1 (5 min, Straße 1, VP 1: Version A, VP 2: Version B)
- Freies Gespräch (15 min)
- Diapix 2 (5 min, Strand 1, VP 1: Version A, VP 2: Version B)
- VP 2 Wortliste 2 (5 min)
- VP 1 Wortliste 2 (5 min)

### 1.3 Material

#### 1.3.1 Freies Gespräch

- Anweisung durch den Experimentator:

„Sie dürfen sich jetzt 15 min frei unterhalten, über beliebige Themen. Ich werde mich nicht einmischen, außer Sie möchten das Experiment abbrechen. Ich gebe Ihnen gerne eine kleine Starthilfe: Wie bewerten Sie denn das Essen in der Mensa?“

#### 1.3.2 Diapixe

Die Diapixe (*Street 1, Farm 1*, Baker & Hazan 2011) für die aufgabenbasierten Dialoge wurden von Alina Zöllner ins Deutsche übersetzt. Eine Ansicht der Originale findet sich hier: <https://www.phon.ucl.ac.uk/project/kidLUCID/diapix.php>, besucht am 07.04.2020.

#### 1.3.3 Wortlisten

Tabelle 3 enthält die Wörter, deren Vokale bspw. als Referenzvokale dienen können. Zusätzlich wurden folgende Wörter gelesen: *ähnlich, Äther, emsig, Ämter*.

**Tabelle 3:** Wortlisten. Alle Wörter sind in den Kontext „Sage X bitte“ eingebettet.

Onsetartikulationsort der Ultima	bilabial		palatal		velar	
Final mit [ə]	piepe	[i:]	Güte	[y:]	bücke	[ʏ]
	Bube	[u:]	Beete	[e:]	Böcke	[œ]
	Kippe	[ɪ]	böte	[ø:]	Pocke	[ɔ]
	Puppe	[ʊ]	Bote	[o:]	Tage	[a:]
			bäte	[ɛ:]	packe	[a]
			Kette	[ɛ]		
Final mit [ɐ]	Pieper	[i:]	Güter	[y:]	Bäcker	[ɛ]
	Tupper	[ʊ]	Puder	[u:]	Höcker	[œ]
	Kaper	[a:]	Mütter	[ʏ]	Packer	[a]
	Geber	[e:]	Köder	[ø:]	Kicker	[ɪ]
			Toter	[o:]		
			Täter	[ɛ:]		
			Dotter	[ɔ]		

## 1.4 Dateienbenennung

Die Daten in BeDiaCo v.1 enthalten zuerst den Registertyp (*frei*, *wortliste*), anschließend bei Wortlisten eine, bei Dialogen beide Versuchspersonen. In Dialogen entspricht der erstgenannten Versuchsperson homolog immer Kanal 1, der zweitgenannten Kanal 2. Welcher Kanal enthalten ist wird durch *ch1/ch2* markiert. Register, Versuchsperson und Kanal werden durch Unterstriche abgetrennt; bei Wortlisten zudem die Nummer der Liste.

Beispiele:

- *frei\_f7m1\_ch1.wav* bezeichnet das Audiosignal des freien Dialogs von f7 mit m1, das Signal selbst enthält nur Kanal 1, in dem Fall also f7.
- *wortliste\_m1\_2.TextGrid* bezeichnet das TextGrid der zweiten Wortliste von m1.

## 1.5 Audionachverarbeitung und Pseudonymisierung

Die Kommunikation zwischen Experimentator und Versuchspersonen zu Beginn und Ende des Experiments wurde entfernt.

Ohne Aufnahmefilter geriet eine Störfrequenz um ca. 50 Hz (wahrscheinlich aufgrund der Stromzufuhr) auf die Audiospur. Zur Auspendelung des Oszillogramms auf die Nulllinie wird in Praat (Boersma 2001) der Filter *Reduce noise* angewendet (für jeden Kanal getrennt!):

1. Rauschunterdrückung (Praat *Reduce noise*)
  - Noise time range (s) 0.0 0.0

- Window length (s) 0.025
- Filter frequency range (Hz) 80 10 000
- Smoothing bandwidth (Hz) 40
- Noise reduction (dB): -20
- Noise reduction method: spectral-substraction

## 2. Pseudonymisierung (Belz 2019)

- Zur Pseudonymisierung bzw. Anonymisierung wird zunächst eine entsprechende Annotationsebene erstellt, auf der potenziell identifizierbare Informationen markiert werden (vgl. Abschnitt 3.7). Anschließend wird das Skript zur Pseudonymisierung verwendet (ebd.).

## 1.6 Überblick über die Annotationen

Tabelle 4 fasst die vorhandenen und annotierten Audioaufnahmen zusammen.

**Nota bene: Diapixe sind in v1 noch nicht enthalten!**

**Tabelle 4:** Zusammenfassung der Annotationen.

Version	Register	Ebene	Annotation vorhanden
1	Aufgabenfrei	dipl	ja
1	Aufgabenfrei	fp	ja
1	Aufgabenfrei	phon	ja
1	Aufgabenfrei	dia	ja
1	Aufgabenfrei	ip	ja
1	Aufgabenfrei	pseudo	ja
1	Aufgabenbasiert	dipl	nicht in v1 enthalten
1	Aufgabenbasiert	fp	nicht in v1 enthalten
1	Aufgabenbasiert	phon	nicht in v1 enthalten
1	Aufgabenbasiert	dia	nicht in v1 enthalten
1	Aufgabenbasiert	ip	nicht in v1 enthalten
1	Aufgabenbasiert	pseudo	nicht in v1 enthalten
1	Wortliste	ORT	ja
1	Wortliste	MAU-cor	ja
1	Wortliste	vowel	ja

## 1.7 Verfügbarkeit

BeDiaCo v.1 ist über das Medienrepositorium (<https://rs.cms.hu-berlin.de/phon/>) der HU für die Lehre an der HU sowie für die wissenschaftliche Forschung im Bereich der Linguistik (vgl. Abschnitt 1.10) verfügbar. Für den Zugriff auf die für wissenschaftliche Zwecke lizenzierten Korpora dort können Sie folgende Optionen nutzen.



- Einen Account beantragen. Hierzu schreiben Sie eine Mail an [phonetik-labor.german@hu-berlin.de](mailto:phonetik-labor.german@hu-berlin.de) mit Ihrem Namen, Ihrer Affiliation und Ihrem Forschungszweck.
- Einen Link zum Download erhalten. Hierzu schreiben Sie eine Mail an [phonetik-labor.german@hu-berlin.de](mailto:phonetik-labor.german@hu-berlin.de) mit dem gewünschten Korpus, Ihrem Namen, Ihrer Affiliation und Ihrem Forschungszweck.

## 1.8 Datenstruktur im Medienrepositorium

Das vorliegende Korpus ist im Medienrepositorium als Kollektion verfügbar (für Zugang vgl. Abschnitt 1.7). Audio-, TextGrid- und Dokumentations-Dateien sind alle auf einer Ebene zu finden (das Medienrepositorium erlaubt keine Ordnerstrukturen). Die Emu-Datenbanken der freien Dialoge und der Wortlisten sind zu zip-Dateien zusammengefasst.

## 1.9 Versuchspersonenmetadaten und Dyadenzuordnung

Alle Versuchspersonen sind deutsche Muttersprachlerinnen und Muttersprachler ohne Hörbeeinträchtigung. Tabelle 5 zeigt die erhobenen Metadaten je Versuchsperson.

**Tabelle 5:** Metadaten der Versuchspersonen und ihre Dyaden.

Dyade	VP	Alter	Geschlecht	Händigkeit	Abschluss <sup>a</sup>	Land <sup>b</sup>	Elter 1 <sup>b</sup>	Elter 2 <sup>b</sup>
b	f1	21	f	rechts	Abi	BE	TH	SN
b	f2	22	f	rechts	Abi	NW	NW	NW
e	f7	25	f	rechts	Abi	BE	BE	BE
e	m1	27	m	rechts	HS	NI	NI	NI
g	m4	19	m	rechts	Abi	BE	TH	TH
g	m5	22	m	rechts	Abi	NW	NW	NI
h	m6	31	m	rechts	HS	BE	BE	BE
h	m7	21	m	rechts	Abi	MV	MV	MV
i	f10	23	f	links	Abi	BE	ST	ST
i	m8	19	m	rechts	Abi	BE	MV	BGR
k	m10	29	m	rechts	Abi	MV	MV	MV
k	f12	24	f	rechts	Abi	HE	HE	HE
m	m13	31	m	rechts	HS	HH	HH	HH
m	m14	18	m	rechts	Abi	BE	BE	BB
n	m15	28	m	rechts	HS	BB	BB	TH
n	f13	26	f	rechts	MR	BE	NW	NW

<sup>a</sup> Abi = Abitur, HA = Hochschulabschluss, MR = Mittlere Reife

<sup>b</sup> BGR = Bulgarien, BW = Baden-Württemberg, BY = Bayern, BE = Berlin, BB = Brandenburg, HB = Bremen, HH = Hamburg, HE = Hessen, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, NW = Nordrhein-Westfalen, RP = Rheinland-Pfalz, SL = Saarland, SN = Sachsen, ST = Sachsen-Anhalt, SH = Schleswig-Holstein, TH = Thüringen

## 1.10 Einwilligungserklärungen

Tabelle 6 zeigt die Einwilligungen je Versuchsperson. Die Tabellenüberschriften entsprechen dabei folgendem Wortlaut in den vorgelegten Einwilligungserklärungen, die einzeln ankreuzbar sind.

- Projekt

Text Ich willige ein, dass meine Daten wie in dem Probandeninformationsblatt beschrieben für die phonetische Studie am Lehrstuhl für Phonetik/Phonologie der Humboldt-Universität zu Berlin unter Leitung der oben genannten Studienleiter gespeichert und verarbeitet werden dürfen.

Audio Ich willige ein, dass meine pseudonymisierten Audiodaten wie in dem Probandeninformationsblatt beschrieben für die phonetische Studie am Lehrstuhl für Phonetik/Phonologie der Humboldt-Universität zu Berlin unter Leitung der oben genannten Studienleiter gespeichert und verarbeitet werden dürfen.

- **Erläuterung:** Dieser Punkt erlaubt die Speicherung, Verarbeitung und Verwendung der Daten innerhalb des Projekts BeDiaCo.
- Lehre
  - Text Ich willige ein, dass meine anonymisierten Daten in der Lehre am Institut für deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin gezeigt werden dürfen.
  - Audio Ich willige ein, dass meine pseudonymisierten Audiodaten in der Lehre am Institut für deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin gezeigt werden dürfen.
    - **Erläuterung:** Dieser Punkt erlaubt die Verwendung der Daten in der Lehre.
- Kongresse
  - Text Ich willige ein, dass meine anonymisierten Daten von Mitarbeiterinnen und Mitarbeitern des Instituts für deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin auf wissenschaftlichen Kongressen gezeigt werden dürfen.
  - Audio Ich willige ein, dass meine pseudonymisierten Audiodaten von Mitarbeiterinnen und Mitarbeitern des Instituts für deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin auf wissenschaftlichen Kongressen gezeigt werden dürfen.
    - **Erläuterung:** Dieser Punkt erlaubt die Verwendung der Daten auf wissenschaftlichen Kongressen.
- Dritte
  - Text Ich willige ein, dass meine anonymisierten Daten unter Einhaltung der geltenden Datenschutzbestimmungen zu Forschungszwecken an andere Wissenschaftlerinnen und Wissenschaftler weitergegeben werden dürfen.
    - **Erläuterung:** Dieser Punkt erlaubt die Weitergabe zu Forschungszwecken an projektexterne Wissenschaftler\*innen. Allerdings muss hierzu ein Vertrag mit diesen geschlossen werden.
- CC-BY
  - Text Ich willige ein, dass die anonymisierten Transkriptionsdaten zusammen mit den Metadaten als Teil eines Korpus unter einer CC-BY-3.0-DE-Lizenz veröffentlicht werden dürfen.
  - Audio Ich willige ein, dass meine pseudonymisierten Audiodaten nach Entfernung jeglicher Klarnamen vom Lehrstuhl für Phonetik/Phonologie der

Humboldt-Universität zu Berlin unter Leitung der oben genannten Studienleiter und vorbehaltlich der Genehmigung durch die Datenschutzbeauftragte der Humboldt-Universität zu Berlin unter der Creative-Commons-Lizenz CC BY-NC 3.0 DE veröffentlicht werden dürfen.

- **Erläuterung:** Mit diesem Punkt wurde versucht, die Daten unter CC zu lizenzieren. Aufgrund erheblicher Unklarheit über die datenschutzrechtliche und urheberrechtliche Möglichkeit einer CC-Lizenzierung von Text- und Audiodaten werden die Daten trotz Einwilligung der Versuchspersonen jedoch vorerst nicht lizenziert.

- Repo

Text/Audio Wir bitten Sie um Ihre Einwilligung, dass wir die von Ihnen bei uns erhobenen und anonymisierten/pseudonymisierten Daten an ein Datenzentrum zur Archivierung und weiteren wissenschaftlichen Nutzung für zukünftige themenverwandte Forschungsprojekte übermitteln werden. Das Datenservicezentrum stellt Wissenschaftlerinnen und Wissenschaftlern die anonymisierten/pseudonymisierten Daten für ausschließlich wissenschaftliche Zwecke im Bereich linguistischer Forschung unter Beachtung der jeweils geltenden Datenschutzbestimmungen zur Verfügung. Hierzu bleiben Ihre Daten über das Ende des aktuellen Forschungsprojektes sowie der von der DFG im Sinne guter wissenschaftlicher Praxis geforderten Nachweispflicht von 10 Jahren hinaus entsprechend gespeichert.

- **Erläuterung:** Aufgrund erheblicher Unklarheit über die datenschutzrechtliche und urheberrechtliche Möglichkeit einer CC-Lizenzierung von Text- und Audiodaten wurde im Nachgang der Datenerhebung eine Mail mit der Repo-Einwilligung an die Versuchspersonen verschickt, mit der letztlich 16 VP einverstanden waren, die somit auch in BeDiaCo v.1 eingehen können (vgl. Tabelle 6).

**Tabelle 6:** Einwilligungen je Versuchsperson.

Dyade	VP	Projekt	Lehre	Kongresse	Dritte	CC-BY	Repo	v
b	f1	x	x	x	x	x	x	1
b	f2	x	x	x	x	x	x	1
e	f7	x	x	x	x	x	x	1
e	m1	x	x	x	x	x	x	1
g	m4	x	x	x	x	x	x	1
g	m5	x	x	x	x	x	x	1
h	m6	x	x	x	x	x	x	1
h	m7	x	x	x	x	x	x	1
i	f10	x	x	x	x	x	x	1
i	m8	x	x	x	x	x	x	1
k	m10	x	x	x	x	x	x	1
k	f12	x	x	x	x	x	x	1
m	m13	x	x	x	x	x	x	1
m	m14	x	x	x	x	x	x	1
n	m15	x	x	x	x	x	x	1
n	f13	x	x	x	x	x	x	1

## 2 Annotationsschema

### 2.1 Aufgabenfreie Dialoge

Die aufgabenfreien Dialoge enthalten sieben Annotationsebenen (vgl. Tabelle 7).

**Tabelle 7:** Annotationsebenen der aufgabenfreien Dialoge und ihre Bezugnahme untereinander sowie auf das akustische Signal (AS).

Ebenenname	Enthält	Bezug auf	Alignierung mit
dipl (3.1)	Diplomatische Transliteration	AS	AS
fp (3.2)	Füllpartikeln	dipl	dipl, AS
segm (3.3)	Segmente der Füllpartikeln	fp	fp, AS
phon (3.4)	Phonationsart von Füllpartikeln	segm	segm
ip (3.5)	Intonationsphrasen	dipl	dipl
dia (3.6)	Dialogstruktur	dipl	dipl
pseudo (3.7)	Anonymisierung von Eigennamen	dipl, AS	dipl, AS

### 2.2 Wortlisten

Die Wortlisten enthalten drei Annotationsebenen (vgl. Tabelle 8).

**Tabelle 8:** Annotationsebenen der Wortlisten und ihre Bezugnahme untereinander sowie auf das akustische Signal (AS).

Ebenenname	Enthält	Bezug auf	Alignierung mit
ORT (3.8)	Diplomatische Transliteration	AS	AS
MAU-cor (3.9)	Füllpartikeln	dipl	dipl, AS
vowel (3.10)	Segmente der Füllpartikeln	dipl	AS

## 3 Annotationsebenen

### 3.1 Ebene dipl (Transliteration)

Name	dipl
Beschreibung	Transliteration als Annotation auf dem akustischen Signal
Annotationsart	Spannenannotation auf akustischem Signal
Erstellung	Halbautomatisch (vgl. 3.1.1) mit CAT.exe (Simon Sauer)
Annotationswerte	Offenes Set (vgl. 3.1.2)
Annotator*innen	Malte Belz, Megumi Terada, Alina Zöllner

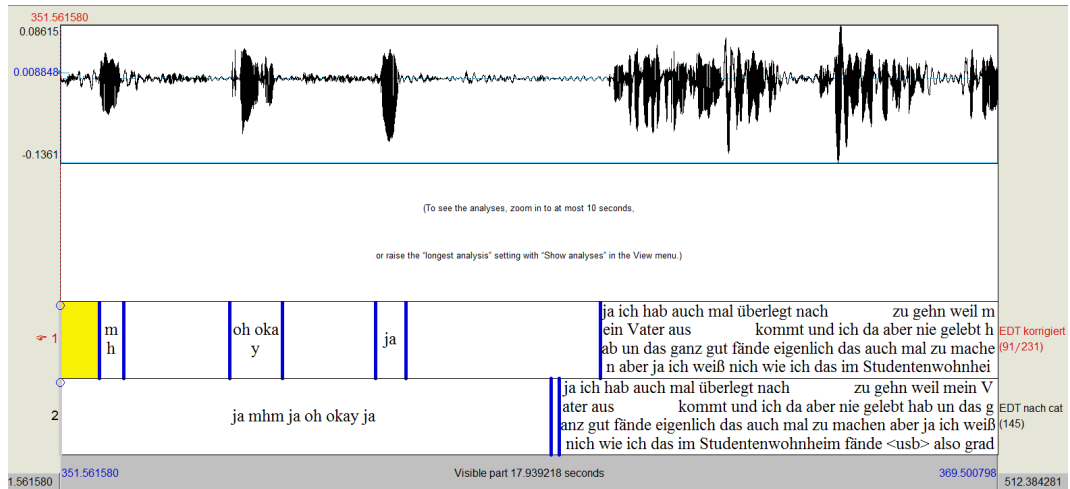
#### 3.1.1 Erstellung

Kurzüberblick:

- Transliteration mit CAT (Sauer o. J.)
- Korrektur der Tokenisierung
- Alignierung mit dem Signal

**Transliteration mit CAT** Vor der Transkription wird die wav-Datei geschnitten, so dass der Experimentator nicht mehr zu hören ist. Die Rauschunterdrückung (Hochpass-Filter, vgl. Abschnitt 1.5) erfolgt **nach** der Transliteration mit CAT (ebd.).

CAT segmentiert die Datei in Chunks, die dann nacheinander transliteriert werden (diplomatische Annotation, vgl. Abschnitt 3.1.2. Extralinguistische Entitäten wie Lachen oder Husten werden als <usb> (*non-understandable word or other human noises*), sonstige Störgeräusche als <nib> (*non-human noise*) markiert. Anschließend wird von CAT eine TextGrid-Datei erstellt, die in Praat (Boersma & Weenink 2019) überprüft wird. Die Grenzen werden so korrigiert, dass **möglichst Äußerungen des/der anderen Sprechers\*in nicht innerhalb der Chunks liegen**. Oft müssen daher Feedback-Einheiten korrigiert werden. Abbildung 1 zeigt die Korrektur der Intervallgrenzen in der Ebene *EDT korrigiert* im Vergleich zur Ebene *EDT nach cat*.



**Abbildung 1:** Vergleich der korrigierten (Ebene 1) und unkorrigierten Chunks (Ebene 2) nach der Transliteration mit CAT und vor der Alignierung mit WebMaus.

Alle überflüssigen Intervallgrenzen wie in Abbildung 2, die von CAT eingefügt wurden, müssen vor der Alignierung mit WebMaus (Kisler et al. 2017) gelöscht werden. Die Grenzen sollten immer möglichst wenig Pausen abbilden. Sehr kleine Pausen (ca. unter 200 ms) sollten gelöscht werden, da WebMaus sonst Fehler ausgibt.

Die mit Cat erstellten und korrigierten TextGrids werden in einem separaten Ordner *cat\_output* gespeichert, der nicht Teil des veröffentlichten Korpus ist.

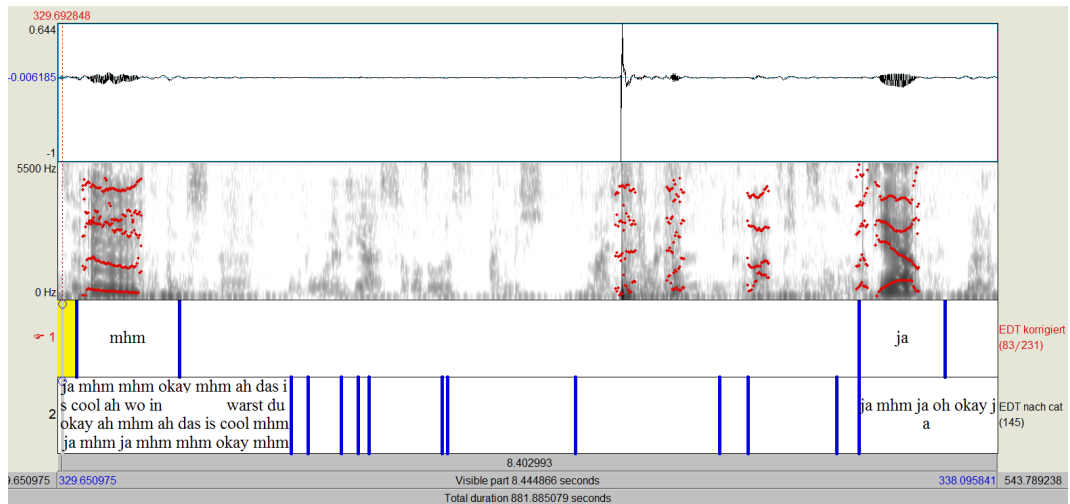
Das TextGrid wird dann mit dem BAS-Service *Chunk Preparation*<sup>2</sup> (Reichel & Kisler 2014) und folgenden Einstellungen in das .par-Format umgewandelt:

- Language: German (DE)
- Input format: tg
- Input tier name: EDT
- Sampling rate: 44100
- Keep annotation: yes

Die Datei mit der Endung .par wird auch unter dem Ordner *dialoge/cat\_output* gespeichert, allerdings genauso benannt wie die wav-Datei, da die Namensnennung für WebMaus gleich sein muss. Beispielsweise ergibt sich dann folgendes Paar, wobei „ch1“ für den ersten Kanal der Dyade steht und hier also Versuchsperson f7 beinhaltet.

- *frei\_f7m1\_ch1.wav*

<sup>2</sup><https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/ChunkPreparation>



**Abbildung 2:** Löschen überflüssiger Intervallgrenzen der Ebene 2 (EDT nach cat).

- *frei\_f7m1\_ch1.par*

Mit *WebMAUS General*<sup>3</sup> (Kisler et al. 2017) wird dann der jeweilige Kanal und die zugehörige par-Datei segmentiert und aligniert, und zwar mit den folgenden Einstellungen (nicht erwähnte Einstellungen bleiben im Default-Modus):

- Input encoding: sampa
- Language: German (DE)
- Inter-word silence: 7
- KAN tier in TextGrid: true
- ORT tier in TextGrid: true
- Chunk segmentation: true
- Pre-segmentation: true
- Output symbols: sampa

Die fertig geMAUSsten Dateien werden mit der zugehörigen wav-Datei unter *3\_dialoge/Buchstabe\_v1* gespeichert. **Buchstabe** steht in diesem Fall für die Session, also beispielsweise *3\_dialoge/e\_v1*.

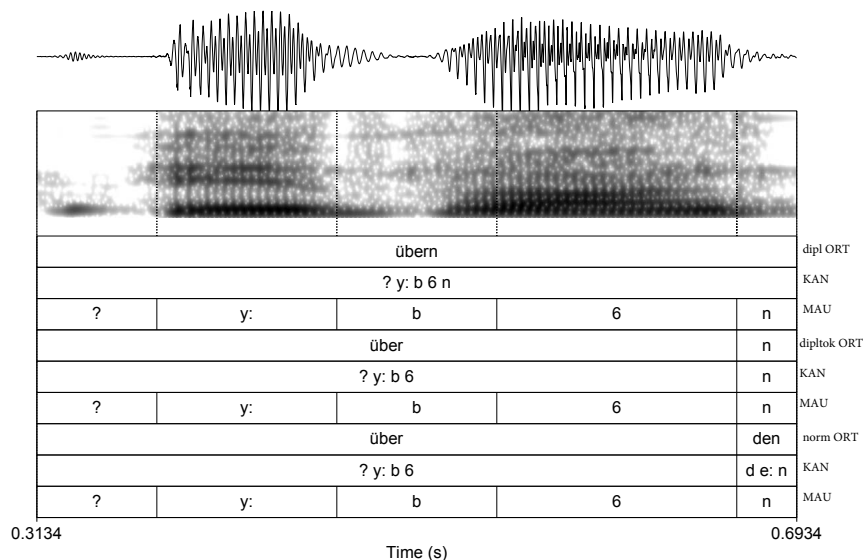
### 3.1.2 Annotationswerte

Aus der Vielfalt der möglichen Transliterationsansätze wird eine diplomatische Transliteration auf Aussprachebasis gewählt. Spontane gesprochene Sprache enthält Eli-

<sup>3</sup><https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/WebMAUSGeneral>



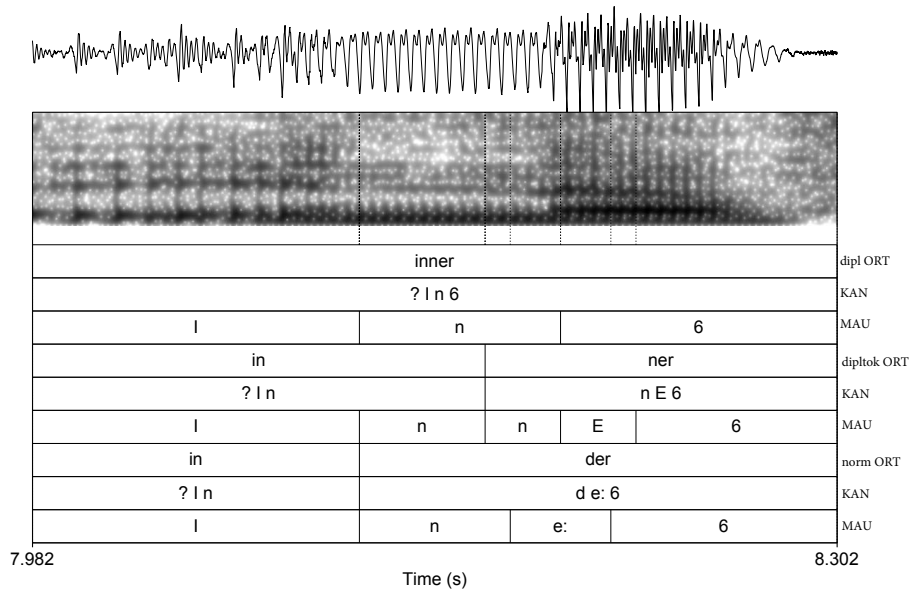
sionen, Epenthesen und Verschmelzungsformen. Während eine von der Schrift abweichende Aussprache von WebMaus zwar erkannt werden kann, ist dies jedoch nicht immer eine gute Wahl. Exemplarisch kann dies mit den Verschmelzungsformen *übern* in *übern Campus* und *inner* in *äh inner Fußgängerzone* gezeigt werden. WebMaus benennt die graphische Transliterationsebene als **ORT**, da hier orthographischer Input erfolgt. In Abbildung 3 und Abbildung 4 werden drei Fälle mit drei verschiedenen Inputsegmentierungen gegenübergestellt: *übern/inner* auf **dipl ORT**, *über n/inner n* auf **dipltok ORT**, und *über den/in der* auf **norm ORT**. **dipl ORT** entspricht einer Art Verschmelzungsschreibweise, wie bspw. *ins*, *fürs*; **dipltok ORT** stellt eine tokenisierte Schreibweise dar, in der der Artikel mit einem Spatium abgetrennt wird, und **norm ORT** stellt eine mögliche standarddeutsche Zielhypothese als Vollform dar.



**Abbildung 3:** Für **dipl ORT**, **dipltok ORT** und **norm ORT** sind die drei von WebMaus zurückgegebenen Ebenen **ORT** (orthographische Transliteration), **KAN** (kanonische Transliteration) und **MAU** (Segmentebene) dargestellt.

WebMaus erkennt für *übern* auf **dipl ORT**, dass hier eine reduzierte Aussprache von *den* vorliegt. In diesem Fall wäre auch eine tokenisierte Transliteration in **dipltok ORT** oder gar eine normalisierte Transliteration wie in **norm ORT** unschädlich für das korrekte Erkennen gewesen. Anhand dieses Beispiels scheint es so, als ob die Auswahl einer Transliterationsweise eine rein theoretisch motivierte sein kann. Nach Betrachtung der Segmentierung verschiedener Transliterationen von *inner/in ner/in der* in der folgenden Abbildung 4 muss dies jedoch revidiert werden.

Das Ergebnis für *inner* auf der normalisierten Ebene **norm ORT** ist problematisch,



**Abbildung 4:** Für dipl ORT, dipltok ORT und norm ORT sind die drei von WebMaus zurückgegebenen Ebenen ORT (orthographische Transliteration), KAN (kanonische Transliteration) und MAU (Segmentebene) dargestellt.

da hier von einer kanonischen Form [de:6] ([de:ɐ]) ausgegangen wird, und das [e:] auf der MAU-Ebene erhalten bleibt. Auch die Segmentierung des Nasals ist sowohl für norm MAU als auch für dipltok MAU eher arbiträr gekürzt. Hier erzielt die dipl ORT-Ebene die beste Segmentierung. Aufgrund dieser Tatsache wird die Transliteration diplomatisch und aussprachebasiert durchgeführt. Mit Hinblick auf Forschungsfragen, welche sich an einer standardnäheren orthographischen Ebene orientieren, kann jederzeit eine normalisierte Transliteration als neue Ebene eingeführt werden.

Tabelle 9 gibt einen nicht-exhaustiven Überblick über die aussprachebasierte Transkriptionen.

**Tabelle 9:** Beispiele aussprachebasierter Transkription, wie sie für BeDiaCo verwendet wird (nicht-exhaustiv).

Standardorthographische Repräsentation	Aussprachebasierte Transkription
in der	inner
nicht	nich
ist	is
haben	ham
aber	a
letztem	letzem
jetzt	jetz
weil es	weils
stelle es	stells
ein	n
interessantes	intressantes
eine	ne
sage	sag
glaube	glaub
anderen	andern
wie es	wies
irgendeinem	irgendeim, irgeneim, ...
denen	den
das	s
machmal	machma
frage	frag
eingentlich	einkich
mal	ma
gerade	grade
andere	andre
könnte	könnt
wäre	wär
glaube	glaub
so ein	son
war es	wars
bekommen	bekomm
und	un
lief es	liefs
gar nicht	garnich
ging es	gings
einen	ein
wie einen	wie n
es wäre	s wär

## 3.2 Ebene fp (Füllpartikeln)

Name	fp
Beschreibung	Annotation von Füllpartikeln und ihrem Mikrokontext
Annotationsart	Spannenannotation
Bezug	dip1 (3.1) und akustisches Signal
Erstellung	Manuell (vgl. 3.2.1)
Annotationswerte	Geschlossenes Set (vgl. 3.2.2)
Annotator*innen	Malte Belz

### 3.2.1 Erstellung

Zur Erstellung wird eine neue Intervall-Ebene in Praat angelegt. Für jeden Kanal wird das Signal manuell angehört und bei positiver Evidenz eine Füllpartikel annotiert.

### 3.2.2 Annotationswerte

Werte	Beschreibung der hierarchisch höheren Werte
f_	Gehört zur Variablen, ist eigenständig abgrenzbar (keine Silbenprolongationen, keine stillen Pausen, keine Atmungspausen).
a_	Antezedens. Linkes Intervall von <i>f</i> . Stille Pausen oder das letzte phonetische Segment des vorangehenden Signals.
p_	Postzedens. Rechtes Intervall von <i>f</i> . Stille Pausen oder das erste phonetische Segment des nachfolgenden Signals.
Werte	Beschreibung aller Werte
fv	Ein vokalisiertes Phänomen ohne jene in <i>fc</i> und <i>fg</i> .
fc	Ein nicht näher spezifizierter Klicklaut.
fg	Eine glottalisierte Phase (kein Vokal, kein Click vorhanden).
fx	Nicht in <i>fv</i> , <i>fc</i> oder <i>fg</i> kategorisierbar.
as	Im Antezedens ist ein <b>Lautsegment</b> enthalten.
ac	Im Antezedens ist ein <b>Click</b> enthalten.
ap	Im Antezedens ist eine <b>stille Pause</b> enthalten.
ar	Im Antezedens ist eine <b>überwiegend stille Pause</b> enthalten, die unbestimmte artikulatorische oder vegetative Reflexe enthält.
ah	Im Antezedens ist eine <b>Atempause</b> enthalten.
ae	Im Antezedens ist <b>Extralinguistisches</b> enthalten (Lachen, Räuspern, Schlucken, Husten, etc.).
at	Im Antezedens ist ein <b>Turn</b> des Dialogpartners enthalten.
ps	Im Postzedens ist ein <b>Lautsegment</b> enthalten.
pc	Im Postzedens ist ein <b>Click</b> enthalten.
pp	Im Postzedens ist eine <b>stille Pause</b> enthalten.
pr	Im Postzedens ist eine <b>überwiegend stille Pause</b> enthalten, die unbestimmte artikulatorische oder vegetative Reflexe enthält.
ph	Im Postzedens ist eine <b>Atempause</b> enthalten.
pe	Im Postzedens ist <b>Extralinguistisches</b> enthalten (Lachen, Räuspern, Schlucken, Husten, etc.).
pt	Im Postzedens ist ein <b>Turn</b> des Dialogpartners enthalten.

### 3.3 Ebene segm (Segmentale Annotation von Füllpartikeln)

Name	<b>segm</b>
Beschreibung	Annotation der Segmente, aus denen die Füllpartikeln bestehen sowie der Segmente, die in den Ante- und Postzedentia enthalten sind.
Annotationsart	Spannenannotation
Bezug	<b>fp</b> (3.2) und akustisches Signal
Erstellung	Manuell (vgl. 3.3.1)
Annotationswerte	Geschlossenes Set (vgl. 3.3.2)
Annotator*innen	Malte Belz

### 3.3.1 Erstellung

Zur Erstellung wird eine neue Intervall-Ebene in Praat angelegt. Die Annotation erfolgt manuell. Die Segmentierung orientiert sich an der Ausdehnung der als unterscheidbar identifizierbaren Laute. Grenzen werden in Praat dort annotiert, wo der Laut gerade schon oder gerade nicht mehr als Laut einer bestimmten Klasse erkennbar ist. Grenzen werden am Nulldurchgang aufsteigender Flanken im Oszillogramm gesetzt.

### 3.3.2 Annotationswerte

Werte	Beschreibung aller Werte
SAMPA	Jedes auf der <b>fp</b> -Ebene mit <i>bs</i> , <i>as</i> , <i>fv</i> , <i>ps</i> , <i>qs</i> annotierte Token wird mit seinem tatsächlichen symbolphonetischen Wert oder Werten (für <i>fv</i> ) in weiter (i. Ggs. zu enger) Transkription mithilfe von SAMPA annotiert (es gelten die Ergänzungen A–F, s. u.). Token, die sich auf <i>fv</i> beziehen, können weiter unterteilt werden (bspw. in zwei Token <i>?</i> und <i>E</i> ).
?, G, GP, GT, GG	<b>Ergänzung A:</b> Ein bis drei glottale Plosive werden mit <i>?</i> annotiert. Sequenzen von mehr als drei glottalen Plosiven, die enger als 50 ms beieinander stehen, werden mit <i>G</i> annotiert. Sequenzen von mehr als zwei glottalen Plosiven, die weiter als 50 ms entfernt auftreten (entspricht 20 Hz), werden mit <i>GP</i> annotiert. Glottale Transitionen antezedenter Segmente in den <i>fv</i> -Vokal werden mit <i>GT</i> annotiert. Die Abgrenzung zum Vokal in <i>fv</i> wird dort getroffen, wo die regelmäßige modale Phonation nach links hin endet. Glottalisierte Phasen mit geschlossenem Mund werden mit <i>GG</i> annotiert. Unter <i>fg</i> wird <i>?</i> annotiert, wenn einzelne Plosive einen Abstand von mehr als 50 ms (entspricht 20 Hz) aufweisen, da diese als Einzellaute wahrgenommen werden. Mit <i>G</i> werden Phasen glottaler Plosive annotiert, die enger zusammenstehen und mehr als zwei sind. Falls Reste eines zugrundeliegenden Sonoranten erkennbar sind, kann dieser annotiert werden (bspw. [u]).
Q	<b>Ergänzung B:</b> Hohe gepresste Stimmlage vor dem Vokal in <i>fv</i> wird auf <b>segm</b> mit Q annotiert.
E	<b>Ergänzung C:</b> Vokale in <i>fv</i> werden immer als <i>E</i> annotiert. Ihre Segmentierung endet nach der letzten erkennbaren Vokalamplitude.
0	<b>Ergänzung D:</b> Mikropausen innerhalb von <i>fv</i> unter 100 ms werden mit <i>0</i> annotiert. Diese können auch das erste Segment unter <i>fv</i> auf <b>segm</b> sein, wenn davor keine Pause vorhanden ist.
#	<b>Ergänzung E:</b> Unvollständige Stimmlippenschwingung mit geringer Amplitude, die an den linken und rechten Enden von Vokalen auftreten kann, wird mit <i>#</i> annotiert.
x	<b>Ergänzung F:</b> Unentscheidbare Intervalle werden mit <i>x</i> annotiert.
—	Token mit <i>ac</i> , <i>pc</i> , <i>ap</i> , <i>pp</i> , <i>ar</i> , <i>pr</i> , <i>ae</i> , <i>pe</i> , <i>at</i> , <i>pt</i> werden mit einem Unterstrich annotiert.
in	Atempausen ( <i>ah</i> , <i>ph</i> ), in denen eingeatmet wird.
ex	Atempausen ( <i>ah</i> , <i>ph</i> ), in denen ausgeatmet wird.

### 3.4 Ebene phon (Phonationsartannotation der Segmente)

Name	phon
Beschreibung	Annotation der Phonationsart
Annotationsart	Spannenannotation
Bezug	segm (3.3) und akustisches Signal
Erstellung	Manuell (vgl. 3.4.1)
Annotationswerte	Geschlossenes Set (vgl. 3.4.2)
Annotator*innen	Malte Belz

#### 3.4.1 Erstellung

Zur Erstellung wird eine neue Intervall-Ebene in Praat angelegt. Die Annotation erfolgt manuell. Es zählt sowohl der perzeptive Eindruck als auch die Unregelmäßigkeit der Perioden im Oszillogramm/Sonagramm (für Glottalisierung) bzw. die Ausprägung des Vokals im Sonagramm (für Behauchung).

#### 3.4.2 Annotationswerte

Werte	Beschreibung aller Werte
glo	Bezogen auf segm für den mit SAMPA annotierten Vokal innerhalb des Intervalls <i>fv</i> von Ebene fp, wenn dieser Vokal <b>vollständig</b> glottalisiert ist.
mod	Bezogen auf segm für den mit SAMPA annotierten Vokal innerhalb des Intervalls <i>fv</i> von Ebene fp, wenn dieser Vokal <b>vollständig</b> mit modaler Phonation artikuliert ist.
asp	Bezogen auf segm für den mit SAMPA annotierten Vokal innerhalb des Intervalls <i>fv</i> von Ebene fp, wenn dieser Vokal <b>vollständig</b> mit behauchter Phonation artikuliert ist.

### 3.5 Ebene ip (Intonationsphrasenannotation)

Name	ip
Beschreibung	Annotation der Intonationsphrasen
Annotationsart	Spannenannotation
Bezug	fp (3.2) und akustisches Signal
Erstellung	Manuell (vgl. 3.5.1)
Annotationswerte	Geschlossenes Set (vgl. 3.5.2)
Annotator*innen	Malte Belz

#### 3.5.1 Erstellung

Zur Erstellung wird eine neue Intervall-Ebene in Praat angelegt. Die Annotation erfolgt manuell.

### 3.5.2 Annotationswerte

Werte	Beschreibung aller Werte
ip	IP. Maximale Ausdehnung einer kohärenten Intonationskontur mit mind. einem nuklearen Phrasenton. Obligatorische Grenzen: Atmungspause, Grenzton Fakultative Grenzen (hierarchische Reihenfolge): segmentale Längung, tonaler Bewegung, Tonhöhenreset, Laryngalisierung, Pausen > 50 ms, syntaktische Merkmale
ipp	IP mit Postposition. Enthält IP mit prosodisch merklich abgesetztem Wort nach dem Grenzton. Eventuell vorhandene stille Pausen links des Wortes dürfen nicht größer als 50 ms sein. Rechts des Wortes folgt eine stille Pause, stille Pause mit phonetischen Partikeln oder eine Atmungspause.
ipx	Unklarheit über die Intonationsphrasenhaftigkeit.

## 3.6 Ebene dia (Dialogzügeannotation in BeDiaCo)

Name	dia
Beschreibung	Annotation der Dialogzüge
Annotationsart	Spannenannotation
Bezug	dip1 (3.1) und akustisches Signal
Erstellung	Manuell (vgl. 3.6.1)
Annotationswerte	Geschlossenes Set (vgl. 3.6.2)
Annotator*innen	Malte Belz

### 3.6.1 Erstellung

Zur Erstellung wird eine neue Intervall-Ebene in Praat angelegt. Die Annotation erfolgt manuell. Für die Annotation wird temporär ein Stereosignal und die dip1-Ebene des\*der Dialogpartners\*in in Praat hinzugefügt. Nach Abschluss der Annotation werden diese Ebenen wieder entfernt.

### 3.6.2 Annotationswerte



Werte	Beschreibung aller Werte
<u>Initiierende Züge</u>	
e	<b>Erzählung.</b> Information, die nicht direkt vom Gesprächspartner eliziert wurde. Tritt häufig konsekutiv nach responsiven Zügen ein.
fe	<b>Entscheidungsfrage.</b> Frage, die mit <i>ja</i> oder <i>nein</i> beantwortbar ist und syntaktisch Verberststellung oder Verbzweitstellung aufweist, additiv mit steigender Intonation. Auch Ein-Wort-Fragen können darunter fallen, wie <i>echt?</i> . Lexikalisch bspw. <i>bist du...</i> , <i>hast du...</i> , <i>VERBst du...</i> , <i>... gell?</i> , <i>...oder?</i> , <i>absichtlich?</i> , <i>gab es...</i> , <i>kann das sein?</i>
fw	<b>W-Frage.</b> Frage, die entweder mit einem W-Wort beginnt oder bei Ergänzung mit einem W-Wort beginnen könnte. Es wird ein Set aus möglichen Alternativen erfragt, dies kann auch mit einer Präpositionalphrase evoziert werden. Lexikalisch bspw. <i>wer</i> , <i>wie</i> , <i>was</i> , <i>wo</i> , <i>wann</i> , <i>wieso</i> , <i>weshalb</i> , <i>warum</i> , <i>an welcher...</i> , <i>in was...</i>
r	<b>Redebereitschaft.</b> Lexikalisch bspw. <i>okay</i> , <i>gut</i> , <i>hallo</i> , <i>hi</i> .
<u>Responsive Züge</u>	
b	<b>Backchanneling.</b> Rückversicherungssignal; zeigt, dass der Partner gehört oder verstanden wurde. Lexikalisch bspw. <i>ja</i> , <i>gut</i> , <i>das stimmt</i> , <i>das ist das Wichtigste</i> , <i>ok</i> , <i>mhm</i> , <i>ja gut</i> , <i>das ist echt eklig</i> , <i>cool</i> , <i>ach so stimmt</i> , <i>würde mich auch ankotzen</i> , <i>ah</i> , <i>ja ich glaub</i> , <i>hab ich auch</i> , <i>nein</i> , <i>krass</i> , <i>genau</i> , <i>oh voll stressig</i> , <i>echt?</i>
ap	<b>Positive Antwort</b> , auch auf Backchanneling möglich. Mit „ja“ paraphrasierbar. Lexikalisch bspw. <i>ja</i> , <i>mhm</i> , <i>ja ich muss auf jeden Fall jeden Tag da sein</i> , <i>ja also ich habe ihn glaube letzten letztes Wintersemester habe ich ihn noch gesehen</i> .
an	<b>Negative Antwort</b> , auch auf Backchanneling möglich. Mit „nein“ paraphrasierbar. Lexikalisch bspw. <i>nein irgendwie nicht weil ich halt ein richtiges Studium haben wollte</i> , <i>ähm nein hier an der Uni</i> .
au	<b>Unklare Antwort</b> , die Antwort ist weder klar positiv noch negativ oder die Antwort wird nicht gewusst. Lexikalisch bspw. <i>Ich glaube nur einer</i> , <i>Jein</i> , <i>ich weiß nicht</i> , <i>da bin ich mir gerade nicht sicher</i>
aw	<b>Komplexe Antwort.</b> Antwort auf W-Frage, die nicht einfach <i>ja</i> oder <i>nein</i> bedeutet.
x. _ . _	<b>Nicht entscheidbar.</b> Unklare Exponenten werden mit <i>x</i> annotiert, ambige Exponenten mit beiden Werten in alphabetischer Reihenfolge mit Punkt getrennt annotiert, bspw. <i>x.b.e</i>

### 3.7 Ebene pseudo (Anonymisierung)

Name	pseudo
Beschreibung	Anonymisierung
Annotationsart	Spannenannotation
Bezug	dipl (3.1) und akustisches Signal
Erstellung	Manuell (vgl. 3.7.1)
Annotationswerte	Geschlossenes Set (vgl. 3.7.2)
Annotator*innen	Malte Belz

#### 3.7.1 Erstellung

Die Audiodaten werden gemergt (stereo) und manuell durchgehört. Eigennamen und sonstige schutzwürdigen Daten, die zur Identifizierung Betroffener dienlich sein kön-

nen, werden annotiert. Danach wird die Ebene **pseudo** für jeden Kanal getrennt zum TextGrid hinzugefügt.

Nach der Fertigstellung wird das Skript in <https://doi.org/10.18452/20145> (Belz 2019) verwendet, um das Sprachsignal in den annotierten Intervallen mit einem 200 Hz-Ton zu ersetzen.

### 3.7.2 Annotationswerte

Werte	Beschreibung aller Werte
x	Eigennamen und sonstige schutzwürdigen Daten, die zur Identifizierung Betroffener dienlich sein können

## 3.8 Ebene ORT

Name	ORT
Beschreibung	Textalignierung der Wortlistenstimuli
Annotationsart	Spannenannotation
Bezug	Akustisches Signal
Erstellung	Automatisch (vgl. 3.8.1)
Annotationswerte	Offenes Set (vgl. 3.8.2)
Annotator*innen	Malte Belz, Megumi Terada, Sarah Wesolek

### 3.8.1 Erstellung

Die von den Versuchspersonen gelesenen Wortlistenstimuli wurden je Liste an Web-Maus übergeben und mit den Einstellungen in Abschnitt 3.1.1 aligniert.

### 3.8.2 Annotationswerte

Werte	Beschreibung aller Werte
offen	Gelesene Stimuli, neue deutsche Rechtschreibung.

### 3.9 Ebene MAU-cor

Name	MAU-cor
Beschreibung	Korrektur der automatischen Segmentierung aus ORT (3.8)
Annotationsart	Spannenannotation
Bezug	ORT (3.8) und akustisches Signal
Erstellung	Manuell (vgl. 3.9.1)
Annotationswerte	Geschlossenes Set (vgl. 3.9.2)
Annotator*innen	Malte Belz, Megumi Terada, Sarah Wesolek

#### 3.9.1 Erstellung

Die Audiodaten werden manuell durchgehört und auf die korrekte Alignierung mit den TextGrids korrigiert.

#### 3.9.2 Annotationswerte

Werte	Beschreibung aller Werte
SAMPA	SAMPA-Werte.

### 3.10 Ebene vowel

Name	vowel
Beschreibung	Annotation der Voll- und Nebenvokale der Referenzwörter (Tabelle 3)
Annotationsart	Spannenannotation
Bezug	ORT (3.8) und akustisches Signal
Erstellung	Manuell (vgl. 3.10.1)
Annotationswerte	Geschlossenes Set (vgl. 3.10.2)
Annotator*innen	Malte Belz, Megumi Terada, Sarah Wesolek

#### 3.10.1 Erstellung

Die Referenzvokale und Nebenvokale der Wörter aus der Wortliste (vgl. Tabelle 3) werden annotiert. Ebenso die Aussprache von *emsig* als [k] oder [ç].

#### 3.10.2 Annotationswerte

Werte	Beschreibung aller Werte
SAMPA	SAMPA-Werte.

## 4 Emu-Datenbank

### 4.1 Erstellen der emuR-Datenbank

Mit folgenden Konfigurationen wird die EMU-Datenbank (Cassidy & Harrington 2001; Winkelmann et al. 2017, 2018) für die aufgabenfreien Dialoge in R erstellt; vgl. auch das Manual zu EMU-SDMS<sup>4</sup> und die Beschreibung des R-Pakets<sup>5</sup>. Zu vorhandenen Formantkorrekturen in der Datenbank und Änderungen der Visualisierung vgl. Abschnitt 4.2 und Abschnitt 4.3.

Zum Erstellen der EMU-Datenbank wird das Paket *emuR* benötigt (Winkelmann et al. 2018). Teile der in diesem Skript vorhandenen Kommentierungen sind von Raphael Winkelmann.

```
#### Dialogdatenbank erstellen ####
#### Build dialogue database ####
library(emuR)
packageVersion("emuR")

## [1] '1.1.2'

path2tgCol = paste0(getwd(), "/data/dialogues")
dir.exists(path2tgCol)

## [1] TRUE
```

TextGrids und Audiodaten liegen zusammen in einem Ordner (hier `corpus/emu`).

```
#### convert TextGridCollection to emuDB ####
convert_TextGridCollection(path2tgCol,
                           dbName = "bediaco_v1",
                           targetDir = paste0(getwd(), "/corpus/emu"),
                           tierNames = c("dipl",
                                          "fp",
                                          "segm",
                                          "phon",
                                          "ip",
                                          "dia",
                                          "pseudo"))
```

---

<sup>4</sup>Manual: <https://ips-lmu.github.io/The-EMU-SDMS-Manual/>, besucht am 07.04.2020.

<sup>5</sup>R-Paket: <https://cran.r-project.org/web/packages/emuR/emuR.pdf>, besucht am 07.04.2020.

Laden der neuen Datenbank und hinzufügen von Link-Definitionen zwischen den Ebenen.

```
#### load database ####
bdc = load_emuDB(paste0(getwd(), "/corpus/emu/", "bediaco_v1_emuDB"))

#### create linkDefinitions ####
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "dipl",
                  sublevelName = "fp")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "fp",
                  sublevelName = "segm")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "fp",
                  sublevelName = "dia")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "segm",
                  sublevelName = "phon")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "dia",
                  sublevelName = "dipl")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "dia",
                  sublevelName = "fp")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "dia",
                  sublevelName = "ip")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "ip",
                  sublevelName = "dipl")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "ip",
                  sublevelName = "fp")
add_linkDefinition(bdc, type = "ONE_TO_MANY",
                  superlevelName = "pseudo",
                  sublevelName = "fp")
```

Jetzt werden die Ebenen automatisch so verbunden, wie in den Links definiert.

```
#### autobuild (which creates actual links) ####
autobuild_linkFromTimes(bdc, superlevelName = "dipl",
                        sublevelName = "fp",
                        convertSuperlevel = FALSE)
# convertSuperlevel = TRUE will convert a SEGMENT level to
# an ITEM (= timeless) level
autobuild_linkFromTimes(bdc, superlevelName = "fp",
                        sublevelName = "segm",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "segm",
                        sublevelName = "phon",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "dia",
                        sublevelName = "dipl",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "dia",
                        sublevelName = "fp",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "dia",
                        sublevelName = "ip",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "ip",
                        sublevelName = "dipl",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "ip",
                        sublevelName = "fp",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "pseudo",
                        sublevelName = "fp",
                        convertSuperlevel = FALSE)
autobuild_linkFromTimes(bdc, superlevelName = "dia",
                        sublevelName = "fp",
                        convertSuperlevel = FALSE)
```

Hier werden die Formanttracks hinzugefügt, die zuvor mit Praat genderspezifisch berechnet wurden (Winkelmann 2015).

```
#### get wav file paths
wavFilePaths <- list.files(paste0(getwd(), "/corpus/emu/",
                                "bediaco_v1_emuDB"),
```

```

                                pattern = glob2rx("*.wav"),
                                recursive = T, full.names = TRUE)
file.exists(wavFilePaths)

# loop through files and call
# function above then save the AsspDataObject to file
# This function written by Raphael Winkelmann can be found here:
# https://gist.github.com/raphywink/2512752a1efa56951f04)
for(fp in wavFilePaths){
  ado = PraatToFormants2AsspDataObj(fp)
  # use praat to calc. formants + convert result to asspDataObj
  newPath = file.path(paste0(file_path_sans_ext(fp), '.praatFms'))
  print(paste0(fp, ' -> ', newPath))
  write.AsspDataObj(ado, file = newPath)
}

# now we have the Praat estimated formants stored in
# SSFF files in the _bndl folder of our emuDB
# we treat them as regular SSFF files and define a track for them:
add_ssffTrackDefinition(bdc, "FORMANTS", columnName = "fm",
                        fileExtension = "praatFms")

# configure EMU-webApp to display formants
sco = get_signalCanvasesOrder(bdc, "default")
set_signalCanvasesOrder(bdc, "default", c(sco, "FORMANTS"))

```

## 4.2 Enthaltene Signalebenen und Korrekturen

Neben den Audiodateien sind in der emuDB auch Grundfrequenz (.f0-Dateien) und über ein Praat-Skript (Winkelmann 2015) ermittelte Formanten (.praatFms-Dateien) enthalten.

Die Formanttrajektorien für F1, F2 und F3 wurden für alle Vokale von Füllpartikeln (Ebene *segm*) sowie für alle Vokale in den Wörtern der gelesenen Wortliste von Malte Belz händisch korrigiert.

## 4.3 Darstellung im Browser

Die Darstellung im Browser (R-Befehl `serve()`) wird über die Datei `bediaco_v1_DB_config.json` geregelt. Die Darstellung für die emuDB der freien Dialoge und der Wortlisten wurde von Malte Belz mit folgenden Einstellungen konfiguriert, so dass sich

Spektrogramm und Formanttrajektorien überlappen und F1 bis F3 angezeigt wird.

```
"assign": [  
  {  
    "signalCanvasName": "SPEC",  
    "ssffTrackName": "FORMANTS"  
  }  
,  
  "contourLims": [  
    {  
      "ssffTrackName": "FORMANTS",  
      "minContourIdx": 0,  
      "maxContourIdx": 2  
    }  
  ]  
]
```

## Literatur

- Baker, Rachel & Valerie Hazan (2011): DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior research methods* 43 (3). 761–770. DOI: [10.3758/s13428-011-0075-y](https://doi.org/10.3758/s13428-011-0075-y).
- Belz, Malte ([in Vorbereitung]): *Die Phonetik von äh und ähm. Akustische Variation von Füllpartikeln im Deutschen*.
- Belz, Malte (2019): *Pseudonymisierung von Sprachchunks in Audiodaten mit Praat und R (Pseudonymization of speech chunks in audio data using Praat and R)*. Humboldt-Universität zu Berlin. DOI: [10.18452/20145](https://doi.org/10.18452/20145).
- Belz, Malte & Christine Mooshammer (2020): *Berlin Dialogue Corpus (BeDiaCo). Version 1*. Humboldt-Universität zu Berlin. URL: <https://rs.cms.hu-berlin.de/phon>.
- Belz, Malte; Alina Zöllner & Megumi Terada (2020): *Dokumentation und Annotationsrichtlinien für das Korpus BeDiaCo Version 1*. Humboldt-Universität zu Berlin. DOI: [10.18452/21361](https://doi.org/10.18452/21361).
- Boersma, Paul (2001): Praat, a system for doing phonetics by computer. *Glott International* 5 (9). 341–345.
- Boersma, Paul & David Weenink (2019): *Praat. doing phonetics by computer*. URL: <http://www.praat.org/> (besucht am 11.02.2020).
- Cassidy, Steve & Jonathan Harrington (2001): Multi-level annotation in the Emu speech database management system. *Speech Communication* 33 (1-2). 61–77. DOI: [10.1016/S0167-6393\(00\)00069-8](https://doi.org/10.1016/S0167-6393(00)00069-8).



- Kisler, Thomas; Uwe Reichel & Florian Schiel (2017): Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347. DOI: [10.1016/j.csl.2017.01.005](https://doi.org/10.1016/j.csl.2017.01.005).
- Rasskazova, Oxana; Simon Sauer & Christine Mooshammer (2014): *Berlin Dialog Corpus (BeDiaCo). Ein multimodales Korpus für Konvergenz- und Dialogforschung*. CLARIN-D-Workshop Postersession in München.
- Reichel, Uwe D. & Thomas Kisler (2014): Language-independent grapheme-phoneme conversion and word stress assignment as a web service. In: Hoffmann, Rüdiger (Hrsg.): *Elektronische Sprachsignalverarbeitung 2014. Tagungsband der 25. Konferenz Dresden, 26. – 28. März 2014*. 42–49. DOI: [10.5282/ubm/epub.22780](https://doi.org/10.5282/ubm/epub.22780).
- Sauer, Simon (o. J.): *CAT - Chunked Audio Transcription*. Version 3.1.2.
- Winkelmann, Raphael (2015): *PraatToFormants2AsspDataObj.R*. URL: <https://gist.github.com/raphywink/2512752a1efa56951f04> (besucht am 07.03.2017).
- Winkelmann, Raphael; Jonathan Harrington & Klaus Jänsch (2017): EMU-SDMS. Advanced speech database management and analysis in R. *Computer Speech & Language*. DOI: [10.1016/j.csl.2017.01.002](https://doi.org/10.1016/j.csl.2017.01.002).
- Winkelmann, Raphael; Klaus Jaensch; Steve Cassidy & Jonathan Harrington (2018): *emuR. Main Package of the EMU Speech Database Management System*. R package version 1.1.1.